How to Conduct Collaborative Action Research

RICHARD SAGOR

ASCD Association for Supervision and Curriculum Development Alexandria, Virginia USA

4
Data Collection

THE GUIDING PRINCIPLE BEHIND THE DATA COLLECTION PROCESS IS that the information collected should be compelling enough to convince any skeptic. Although the primary audience for action research is the action research team itself, a team sometimes needs to win the support of resident skeptics or cynics in the school; skeptics or cynics usually need overwhelming data to convince them of anything. And in most school districts, conservative pockets in the hierarchy will resist changes in program direction whenever possible. Collecting data that are valid enough to allay concerns (or to at least create cognitive dissonance for the resisters) is the main function of the data collection process.

Three Important Qualities

Researchers have traditionally been concerned with three qualities of their results: validity, reliability, and generalizability. Before we discuss the data collection process any further, let's briefly consider each of these qualities.

Generalizability

Generalizability refers to the applicability of findings to settings and contexts different from the one in which they were originally obtained. Historically, many educational researchers have tried to avoid research that might lead to findings that could be attacked as being too limited in scope or specific only to a certain context. Instead, they have sought to identify factors that apply in many contexts. For instance, the work on teaching effectiveness done by David

Berliner, Madeline Hunter, and others sought to identify those general factors of instruction and lesson design that were applicable across all subjects, grade levels, and categories of students. Similarly, the work on effective schooling sought to find those general characteristics of schools that correlated with effectiveness regardless of the unique attributes of the communities in which they developed.

It is true that much good has come from such work, but its limitations are obvious to any classroom teacher. One reason most teachers are skeptical of this type of research is that they intuitively sense the powerful pull of context. After all, teachers know that what worked with their third period class may well fail with their sixth period class, even though both groups of students are using the same materials. If an idea that worked with one group of students at 10:00 a.m. can't be guaranteed to work with another group of students from the same school at 2:00 p.m., then classroom context must be an important factor. Action researchers certainly think so. Instead of attempting to compensate for context, as scientific researchers do, action researchers embrace context as an integral part of their work.

As action researchers, you don't need to worry about the generalizability of your data because you are not seeking to define the ultimate truths of learning theory. Your goal is to understand what is happening in your school or classroom and to determine what might improve things in that context. Although it is important for you to understand and explain the intricacies of the local context, it is not your mission to declare your context as either typical or atypical. That is a decision to be made by others who observe your work or read about it. For example, in his book *Among Schoolchildren,* Tracy Kidder (1989) does not attempt to argue that "Ms. Jones" is a typical 2nd grade teacher. Instead, he leaves it to his readers to decide whether the teacher's experiences relate to their situations. Likewise, individual readers must decide whether the suffering of Hamlet is unique to his circumstances or is shared by the whole of the human race. Shakespeare, like the action researcher, leaves judgments about generalizability up to the reader.

The other two quality issues involving your results,

validity and reliability, are as applicable to collaborative action research as they are to traditional research studies.

Validity

The concept of validity is found in the question "Do these data really measure or represent what we say they do?" This is a crucial issue. Superstition, bad practice, and even malpractice could be the result of action taken on the basis of invalid data. In *The Mismeasure of Man,* Stephen Jay Gould (1981) points out that not long ago, the pseudo-science of phrenology was accepted as valid practice. Phrenologists held that the size and shape of a person's skull were an indicator of the intellectual capacity and character of that person. If our schools were still influenced by the invalid assumptions of the phrenologists, students with large hat sizes would be assigned to the gifted program, while students with small hat sizes would be assigned to special education courses. Such an outcome would be humorous were it not for the fact that it actually happened. The consequences for students are real and severe when important educational decisions are based on invalid instrumentation, data, or theories. For this reason, as action researchers, we should always strive to be able to defend our findings and reports as valid, meaning they measure what they claim to measure.

Reliability

Reliability concerns the accuracy of your methods of measurement. If each subsequent administration of a particular instrument produces wildly different results, then the instrument is deemed to be unreliable. For instance, as teachers, we are familiar with the practice of using standardized achievement tests to reliably measure the performance of a large group of students. The aggregate score posted by a group of students at a large school on a test given last Tuesday won't differ significantly from the aggregate score posted by the same students on the same test given the following Wednesday. Standardized tests are, however, grossly unreliable in their characterization of the abilities of *individual* students. On most norm-referenced tests, a student's score could differ by as much as 30

percentile points from one administration to the next. The difference here is so extreme that a student could be labeled gifted one week and learning-disabled the next. Clearly, the consequences of making programmatic decisions based on unreliable data (i.e., data that are the results of an instrument with a high margin of error) are quite serious.

Action research teams should begin their planning for data collection guided by three questions:

- Do the instruments and methods we plan to use measure what we claim they do?
- Do the instruments and methods we plan to use accurately measure the phenomena we are studying?
- Will a skeptic be convinced by the weight of the data we amass?

We will be returning to how collaborative action researchers satisfy themselves on these three fronts later in this chapter. Let's look now at some of the data collection options available to your team.

Sources of Data

One of the factors that has limited the collection of data for school and classroom decision making is the belief that good data are both difficult and expensive to collect. Many of us have been led to believe that only data obtained with highly technical instrumentation can yield valid results. The experience of hundreds of action researchers, however, has demonstrated that these beliefs do not reflect reality; schools are naturally data-rich environments, and simply opening our eyes to some of the most frequently used and easily obtainable sources of data can make planning the data collection process much easier. Remember that there is no one correct source of data for school-based action research. The best source will be determined by the research questions your team is asking. Here are some of the more common sources of data:

- *Existing Sources*
  - Student work (portfolios)
  - Archival evidence

- *Tools for Capturing Everyday Life*
  - Diaries, logs, journals
  - Videos
  - Photographs
  - Shadowing
  - Observational checklists/rating scales
- *Tools for Questioning*
  - Interviews
  - Written Surveys
  - Tests

As you can see, the data sources are listed under three categories. The items in the first category are data that already exist in every school and classroom. To use this information in your action research, all you need is permission to use it, a way to access it and save it, and a way to sort it that will give it meaning. The items in the second category are all tools for capturing the richness and texture of the everyday life of schools and classrooms. Finally, the items in the third category are questioning tools that will require the development or acquisition of instrumentation. I will discuss each of these three categories separately and elaborate on a few of the more commonly used sources of data.

Existing Sources

The files in the office of every school in America contain data on student attendance, family background, standardized achievement scores, disciplinary referrals, grades, and previous schools attended for nearly all students. These files can often provide answers to our questions about patterns in the behavior and academic performance of students. Getting at the data may take time, and organizing such information is often hard work. Administrators can provide valuable support for school-based action research teams by supplying clerical help to assemble and sort the data.

In one school where I worked, we conducted an academic post mortem of a class that had graduated the previous spring. We first divided the class into four achievement groups (based upon the quartile of their 9th grade standardized achievement tests). We then proceeded to compare these four groups based on gender, courses taken, special services used, number of years in the system, grades earned, extracurricular activities, disciplinary incidents, attendance, post–high school work or education placement, and 12th grade standardized achievement scores. From this study, we developed a clear picture of the academic careers of four (and when divided by gender, eight) different categories of students. We were able to understood more about our grading patterns, our disciplinary practices, the success or lack of success of our Chapter I and special education programs, the reach of our activities and athletic programs, and our grouping practices. Remarkably, all of this was accomplished without contacting a single student; all the data we needed were in the files, thus, a few hours of secretarial time provided the faculty with mounds of useful data on which we could base our school improvement efforts.

Another type of readily available data is students' work. The value of maintaining student portfolios has been discussed by many educators, and the practice has much to commend it. One of the side benefits of the use of portfolios is that it gives teacher-researchers ready access to meaningful longitudinal data on student growth and development. Anyone who has viewed old home movies or paged through a well-maintained baby album can understand the significance of being able to examine various stages of development. When teachers want to quantify the results of student growth, they can simply ask teachers of similar subjects and grade levels to score pieces of students' work done throughout the semester or year. The identity of students is usually kept from the scorers, and each student work is scored by two or three teachers to ensure the reliability of the findings.

Tools for Capturing Everyday Life

Often the questions or issues for which we need to acquire data call for us to take a critical look at exactly how we are spending our time. But how do we find ways to step back and look objectively at ourselves? Many people believe that this outside-looking-in perspective can be obtained only by using an outside investigator. Not so! Action researchers have within themselves many of the

means necessary to take a fresh look at the world they are immersed in. Here are some examples:

Journals. Suppose that twenty teachers in an elementary school have decided to implement whole-language instruction. The faculty wants to examine the implementation issues involved in establishing this program. Now imagine that every teacher agrees to spend only ten minutes every day (immediately after the children leave) for three weeks writing in a journal responses to this prompt: "Describe what went well, what went poorly, or what was a surprise for you today during language arts instruction."

It's unlikely that anyone would be overburdened by an additional ten minutes per day given to personal reflection, yet in just two weeks the teachers would have 280 separate journal entries that could be discussed and analyzed by the team. This is an example of a small investment producing a treasure trove for an action research team.

Videotape and Photography. Videotape and photography may at first strike you as being too subjective a way to capture the reality of daily life. If you reflect further, however, I think you'll find that this is not necessarily so. If you have ever looked at books of photo essays about countries, regions, or national parks or paged through a truly inspired high school yearbook, you have probably said to yourself, "Yes, that *is* what it's like." Well-chosen photos and videos can portray the textures and realities of schools and classrooms in a vivid light. Conversely, they can also be used to distort reality. Later in this chapter, however, we will discuss triangulation, which is the action researcher's chief protection against allowing one distortion to slant an entire research project.

Shadowing. Shadowing is a tool that can help us see people and circumstances as they really are. It gives us an opportunity to experience a situation in a most naturalistic manner. With school-based action research, the shadowing process involves following selected students or teachers for a specified period of time (generally a day) to collect a picture of a typical day in their life. At a high school where I worked, we conducted such self-examinations annually. Twelve teachers shadowed twelve students (three from each grade level, including a high-, middle-, and low-

achieving student from each grade) for one day each. At the end of the twelve-day shadowing period, the students and their shadows met to compare notes. Ultimately, the twenty-four participants (shadows and shadowees) met with the entire faculty to discuss similarities and differences in the way students of different ages and performance levels experienced our school. (A more detailed description of this process can be found in Sagor 1981.)

Checklists and Rating Scales. Checklists and rating scales require less work than shadowing, yet they can still yield rich data for analysis. The checklist/rating scale process involves three steps:

1. Agree on what constitutes a particular behavior or category.
2. Visit classrooms to collect data.
3. Arrange the data into categories.

With the checklist process, one or two people can collect all the data, thereby saving time for the team, yet enabling everyone to have faith in the data. The key is to have prior agreement on the criteria used for rating. Figures 4.1 and 4.2 on the next two pages show data obtained through checklists from an observer who took readings on student and teacher behavior every thirty seconds. This research was conducted to assess the schoolwide use of class time by students and teachers.

Figure 4.3 on page 38 reflects the results of a series of classroom observations lasting ten minutes each. The observers focused on trying to categorize the cognitive level of the instructional activities pursued.

These uses of checklists provided vivid portraits of how class time and activities were being used. Although the data collection involved only one collector, the data were available to everyone in the action research teams.

Tools for Questioning

Often the information we are looking for isn't immediately available through simple observation, so we have to delve into what people know, believe, or feel about an issue. In such cases, we are inclined to collect data through the use of tests, surveys, and interviews.

Tests. Tests are generally used to ascertain levels of

Figure 4.1
632 Snapshots of Teacher Activities

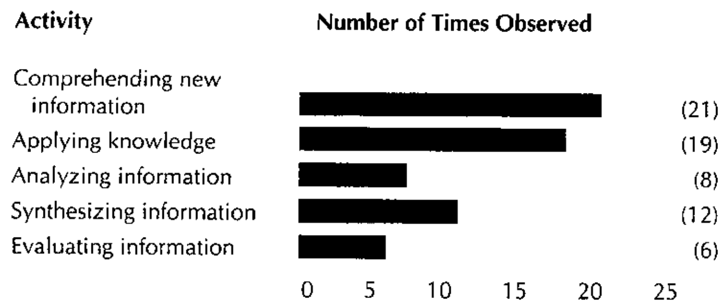
Activity	No. of snapshots where observed	Percentage of total snapshots*
Responding to or helping individual students	112	17%
Asking questions	100	15%
Lecturing	91	14%
Paperwork (at desk)	42	7%
Active monitoring (looking for students needing help)	41	6%
Getting materials	38	6%
Administering tests (proctoring)	27	4%
Management (attendance, organizational details)	27	4%
Audio-visual equipment set-up	25	4%
Observing student performance	23	4%
Responding to student questions	20	3%
Absent from room	16	3%
Giving directions on an assignment	12	2%
Demonstrating for students	11	2%
Cleaning up after lab	9	1%
Small-group instruction	8	1%
Correcting homework	5	1%
Listening to recitations	5	1%
Miscellaneous	5	1%
Admonishing student(s)	5	1%
Waiting for attention	2	.3%
Monitoring student behavior	2	.3%
Calling on a student	2	.3%
Discussion	2	.3%
Dealing with interruptions	1	.2%
Social Interaction with students	1	.2%

**Note:* Due to rounding off, these percentages do not total 100 percent.

Figure 4.2
643 Snapshots of Student Activities

Activity	No. of snapshots where observed	Percentage of total snapshots
Seatwork	142	22%
Listening to teacher	112	17%
Responding to questions from teacher	91	14%
Lab work	88	14%
Taking tests	65	10%
Visiting with each other	25	4%
Performing	18	3%
Reciting	14	2%
Asking the teacher questions	14	2%
Watching or listening to to audio-visual materials	13	2%
Getting out materials	12	2%
Group work	10	2%
Clean-up	10	2%
Observing performance	8	1%
Discussion	8	1%
Responding to questions in writing	5	.8%
Responding to management questions	4	.6%
Observing teacher demonstrations	2	.3%
Waiting	2	.3%

Figure 4.3
Types of Cognitive Activities Observed in a Series of Classroom Observations



skill and knowledge. Action research teams can take advantage of already-developed tests, norm-referenced standardized tests, criterion-referenced tests, Regents' exams, or tests developed by textbook publishers. Occasionally teams may want to develop tests of their own. Because of the importance of obtaining valid, reliable data, it is generally a good idea to consult with a critical friend before administering a test as part of an action research project (the role of the critical friend will be discussed at the end of this chapter).

Written Surveys. Written surveys are favorite data-gathering techniques for action researchers because they enable us to get at emotional, cognitive, and attitudinal issues. Here are some guidelines for developing surveys:

- Surveys can be used to gather data concerning affective, cognitive, or attitudinal issues, depending on how we frame the questions. For example, we might ask questions like:
  a. What are our current school goals? (cognitive)
  b. What is your opinion of the quality of our school goals? (attitudinal)
  c. How do you feel about working on our school goals? (affective)

- Strive for clarity. Conduct a small field test to make sure participants will understand your survey questions.
- Strive for brevity. Surveys should take no longer than ten minutes to administer.
- Provide an opportunity for participants to make additional comments.
- Consider disaggregation and confidentiality issues in advance.
- Remember that your first goal is face validity. Will a skeptic believe your results?
- Avoid suggesting a desired response (asking "leading questions"). For example, instead of asking, "What is the effect of tardiness on achievement?" ask, "What factors influence achievement in our school?" or "What is the effect of tardiness in our school?"
- Try to separate fact from opinion by asking follow-up questions like "How do you know? or "What led you to this opinion?"
- When using numerical scales, ask for an explanation of responses. For example:

How would you rate discipline in this school?

Poor		Average		Excellent
1	2	3	4	5

What would it take to make your rating a five?

- When using numerical scales, consider whether you want a scale that forces people to respond either positively or negatively (an even-numbered scale) or a scale that allows people to respond neutrally (an odd-numbered scale).
- Finally, remember that when you collect data from people, you must be willing to share with them your results.

Interviews. When using interviews to collect data, it is helpful to work from an interview guide, which is simply a general road map to help the interviewer direct the discussion in a manner that is likely to produce relevant information and insights. Although an interview usually requires considerable time commitments from both interviewer and interviewee, it is an excellent way to collect data. Most people enjoy being interviewed because your interest tells them you really care what they think;

nevertheless, their time must be respected. For this reason, you should avoid the need for follow-up interviews. You can do so by using the interview guide to keep yourself on track. Here are some basic points to consider when developing your guide (note that these guidelines are very similar to those for developing surveys):

- Limit yourself to ten to fifteen question areas.
- Reread the research questions the team developed in the problem formulation stage before writing the first draft of the guide.
- Review the research questions after writing the guide to be sure you have asked for all relevant information.
- Follow all your factual and opinion questions with "why" probes.
- Decide how the interview data will be collected: through videotape or audiotape, your notes, or the notes of a co-interviewer.
- Remember that interviews may be used to obtain cognitive, opinion, and affective data. For example, you might ask:
  - What do you recall about our current school goals? (cognitive)
  - What is your opinion of the quality of our school goals? (opinion)
  - How do you feel about working on our school goals? (affective)
- Strive for clarity. Practice your interview with a friend.
- Strive for brevity. An interview should last no longer than forty-five minutes.
- Provide an opportunity for respondents to give full explanations.
- Consider disaggregation and confidentiality issues in advance.
- Avoid suggesting a desired response through a leading question. For example, instead of asking, "What is the effect of tardiness on achievement?" ask "What factors influence achievement in our school?" or "What is the effect of tardiness in our school?"
- Try to separate fact from opinion by asking follow-up questions like these: "How do you know?" "What led you to this opinion?"

- When using numerical scales, always ask for an explanation of responses. For example:

How would you rate discipline in this school?

Poor		Average		Excellent
1	2	3	4	5

What would it take to make your rating a five?

The best aspect of interviews and surveys is that they provide us with actual voices and precise words. Later, when we report our data, this type of information can add credibility to our reports and vitality to our findings. On occasion, however, action researchers have found themselves in trouble by not considering a few basic issues before using these questioning methods. The three most important issues to consider before embarking on data solicitation are confidentiality, disaggregation, and reporting back.

Confidentiality/anonymity. One reason for collecting data through open-ended surveys or interviews is to capture the precise words of your subjects. Before you collect this information, however, you will need to have thought through exactly how you intend to make use of the data obtained. Do you intend to quote sources by name or will you hide their identity? Is it possible to shield identities, or will a careful reader be able to puncture your shroud of secrecy? At the start of any interview or survey, you will want to tell your subject(s) how the data will be used and what confidentiality (if any) will be maintained. Having given these assurances, you *must* maintain them, for your integrity and the opportunities for other action researchers following you depend on your doing so.

A second and related issue involving surveys is whether the individual subjects will be (or need to be) anonymous to the researchers. There is much debate on this topic. Some researchers believe that the freedom to speak anonymously encourages honest and open responses. Others feel that it invites irresponsible and unaccountable cheap shots. You will want to discuss and consider this issue before administering any survey.

Disaggregation. When you are reporting on your data, you may wish to total and analyze responses based on the subpopulations from which they were derived. For example, did the opinions of the girls in the classes differ from the opinions of the boys? Did the outlook of teachers who had been in the building for more than ten years differ from that of new teachers? Did parents with higher levels of education perceive our school more or less favorably than parents with less formal education? If you decide that you want to draw inferences like these from your data (survey and interview data are particularly good for this purpose), you will also need to obtain appropriate demographic data through your survey or interview. Be careful that your requests for demographic information don't compromise any assurances of anonymity. For example, if you ask teachers in a small school to tell you their gender, number of years in the school, and teaching assignment, they may rightly believe that anonymity is impossible.

Reporting back. As mentioned earlier, people generally are flattered to have their opinions solicited and will be generous with their time; they are justified, however, in wanting something in return. What they usually want is feedback. At the time of your survey or interview, you need to be prepared to tell your subjects how you intend to report back any results obtained. If respondents believe you are sitting on unflattering data, your future credibility will be irrevocably damaged.

Capturing the Data. Verbatim responses provide one of the richest sources of information you will find. When a research report quotes the responses of individuals, the report becomes the equivalent of a geography textbook filled with maps, photos, and colorful illustrations. To ensure that you do have verbatim quotes, it is best to tape-record interviews (always with the permission of the person you're interviewing). Tape-recording allows you, the interviewer, to listen actively and give your subject both eye contact and your undivided attention. It also allows you to have a complete and accurate record of the discussion. If tape-recording is impossible, you may want to try using two interviewers; one person can be the listener while the other tries to take verbatim notes.

Triangulation

Earlier in this chapter, we discussed the importance of validity and reliability. Ensuring that research findings are of high quality is as important an issue for the action researcher as for anyone doing scientific inquiry, for conducting research is a time-consuming activity, and you don't want to waste your time collecting unreliable data. Unreliable data will cause you to lose confidence in your results and will cause others to question the validity of your conclusions. Remember that the purpose of collaborative action research is to help you *take action* to improve your school or classroom. If the changes you undertake are based on ill-formed ideas that are the result of invalid and unreliable data, they will probably fail. And the ultimate result will be students who don't perform as well as they could and teachers who feel less competent. Sound research, on the other hand, usually leads to equally sound recommendations for change that will improve student performance and enhance teachers' professional self-esteem.

The quality issue raises a perplexing question for action researchers. Short of investing the time and energy required to become experts in research methodology and instrument development, how can teacher-researchers be sure that their data collection methods are reliable? Realistically, there is no sure solution here, but paying attention to a technique called triangulation will go a long way toward improving the quality of your findings. Simply put, triangulation involves collecting multiple sources of data for every phenomenon or issue being studied.

In our work with collaborative action researchers, we have operationally defined triangulation as using at least three independent windows on whatever phenomenon is being studied. How does this help ensure quality? Let's look at an analogy. Imagine a 3rd grade science lesson on the subject of "observation." The teacher asks Mary to look through the front panel of the classroom terrarium and list or draw everything she sees. Mary, being a diligent student, writes up a thorough list. She's just about to go back to her seat when the teacher asks her to take a look through the side panel of the terrarium. Mary does, and immediately

sees several plants and animals that had been obscured from view in the front panel by rocks and shrubs. By using this second "window" on the phenomenon (the terrarium), Mary now has a more complete picture.

Although Mary feels that her work is now done, the teacher makes one last request. She asks Mary to peer through the top of the terrarium to see if there is anything else she has missed. Mary adds to her list and then sits down. Although Mary may still have failed to observe something of importance, her three windows reveal a far more comprehensive picture than any one window alone could have.

Let's look at another example of triangulation, this one involving an investigation of classroom practice. Say I wanted to investigate my use of cooperative learning structures in my classroom. I might choose to have a colleague observe my class, I might evaluate my own performance as captured on videotape, and I might have another colleague interview my students. If all three windows on my cooperative learning lesson end up showing the same picture, then that picture is likely a valid portrait of my teaching. But suppose that I rated my instruction as inspiring, while my colleague and my students saw it in quite a different light. That discrepancy (a potential source for further inquiry) would have been lost had I used a single measure.

I recall that one school we worked with was investigating the causes of student alienation. The action researchers surveyed students, parents, and teachers. On the issue of staff caring the teachers gave themselves high marks (an average of 9 on a 10-point scale), while students and parents viewed teachers as aloof and cold. More particularly, the less successful students felt strongly that their teachers didn't care about them. Although this discordant information was painful for the teachers to consider, it became the inspiration for much valuable work. Teachers became intensely interested in finding out exactly what they were doing that caused students and parents to perceive them this way. The response to this one item on the survey was a primary force behind the introduction of a teacher advisory program in the school.

In summary, triangulation provides these benefits:

- It compensates for the imperfections of data-gathering instruments.
- When multiple measures yield the same results, it can increase confidence in the results.
- When multiple measures fail to yield the same results, it can raise important follow-up questions.

Making a Data-Collection Plan

Using a matrix like the one in Figure 4.4, you can develop a complete data-collection plan for your project. First, list each of the research questions that surfaced in the problem formulation process. Then for each question, determine at least three independent windows for collecting data on the question being investigated.

Figure 4.4
Data Collection Plan

Research Questions	Data Source #1	Data Source #2	Data Source #3